Because it's there! Why some children count, rather than infer numerical relationships

Kevin Muldoon
Charlie Lewis
John Towse

Lancaster University

Key Terms: Sharing, counting, one-to-one correspondence, numerical equivalence

---

Address for Correspondence:

Kevin Muldoon, School of Life Sciences, Heriot-Watt University,

Edinburgh, EH14 4AS, UK.

Email: k.muldoon@hw.ac.uk Phone (UK) 0131 4518002 Fax (UK) 0131 4513009

Acknowledgements: We would like to thank the staff and children of Greaves Nursery, Lancaster and Morecambe College Playgroup, and St. Martin's Preschool Centre, Lancaster, for their generosity and commitment to this project. The first-named author also wishes to thank the Economic and Social Research Council, UK, for funding a post-doctoral fellowship during which this paper was produced.

Abstract

Two experiments are described that investigate the ability to infer the number of items in one-to-one corresponding sets for two age groups. We assess the influence of set-size, the visibility of sets, and the way in which set equivalence is derived - pairing versus sharing - using a repeated-measures design. Three-year-olds are largely restricted to inferring number after separating out conceptually paired items. In contrast, four-year-olds are able to make appropriate inferences about shared items, but they typically prefer to count those items if they are visible. Moreover, the size of corresponding sets affects children's propensity to count rather than infer. Children count more often on larger sets. The ability to infer number using cardinal extension is associated most strongly with sharing proficiency, although counting skills also play an important part. We discuss how the data reveal an emerging understanding of the relationship between one-to-one correspondence and cardinality.

Because it's there! Why children count, rather than infer, corresponding sets.

The development of children's numerical skills appears to incorporate several paradoxes. As Feigenson, Dehaene & Spelke (2004) note, mathematics is at once both obvious and obscure. On the one hand young children show sophisticated insights into the preciseness of numbers (Sarnecka & Gelman, 2004) and infants distinguish both discrete and continuous quantities (see Feigenson et al., 2004). On the other hand, they also reveal some striking limitations, for example failing to discriminate between numerosities (Xu & Spelke, 2000). In the specific domain of children's counting too, there is a conundrum. Even by the age of approximately three years, children appear to possess a command of core principles that underlie counting behaviour (Gelman & Gallistel, 1978; Gelman & Meck, 1983). Yet there is often an exasperating reluctance among children to use the products of counting to draw legitimate and sensible conclusions about the relationship between sets (Michie, 1984a). In this paper, we present data from two experiments that help to understand why preschool children may fail to make the most of the numerical strategies that they implement. For this, we use a sharing task, and in so doing, we address and attempt to reconcile some apparent inconsistencies in the literature.

Sharing, like counting, is a skill that is acquired and practiced during early childhood (Pepper & Hunting, 1998). Preserving numerical correspondence among shares, through a scheme such as "one-for-you, one-for-me", ensures that each recipient has an equivalent set. However, while this is an appropriate procedure to share out equally, and is a common strategy that appears fairly stable between preschool and school years (Miller, 1984), preschool children may have difficulty realizing what they have accomplished. In this sense, procedural mastery outstrips conceptual awareness of what has been achieved. That is, children below eight years of age often insist that counting is necessary to reach a conclusion about the equality of the shares (Davis & Pitkethly, 1990; Frydman & Bryant, 1988). However, (and in contrast with this research) some findings suggest that, with conceptually paired items, three-year-olds are able to recognize that corresponding sets are equivalent (Sophian, 1988; Sophian, Wood, & Vong, 1995). Why might children be inconsistent in these two situations?

Evidence that sharing has been carried out evenly, or fairly, ought to be sufficient for children to conclude that there are as many items in one set as there are in the other(s). Using a 'same/different' task, Cowan and Biddle (1989) found that three-year-olds understood that distributing items using a 'one-for-you, one-for-me' rule results in sets that are the same, particularly if shares were small (i.e., < four items) and visible. They were not tested on whether they knew how many items each set contained, however, and young children experience difficulty mapping the insights from quantitative equivalence onto number words. Frydman and Bryant (1988) exposed this difficulty, when they asked four-year-olds to share out sweets between dolls, then counted aloud one set for children, and asked how many sweets there were in the other. Even though most children had shared out successfully, none would infer (via the principle of cardinal extension) the numerosity of sweets in the second set. Without exception, every child started to count. Even when the set was hidden so as to preclude counting, less than half the children used equivalence to infer set size.

Using cardinal extension to infer equivalent shares may be difficult for children, but they seem to have several conceptual building blocks in place to reach this milestone. First, children can use knowledge of cardinality when setting up corresponding sets (Becker, 1989; Sophian, 1988). Five-year-olds are able to count-out items and stop when they reach the number which ensures that two sets match (Russac, 1978). Secondly, they are also aware that cardinal extension is appropriate when inferring a number of conceptually paired items. Sophian et al. (1995) showed three-year-olds pairs of frogs and boats (i.e., one frog per boat). These pairs were separated, by placing the frogs inside a box (a narrative explained that they were going to a party) while the boats were (moored) outside. The majority of children then successfully inferred the number of hidden frogs after counting the boats, at least so long as small sets (two or three) were used (consistent with Cowan & Biddle, 1989). Thus, many children apparently knew that correspondence between pairs of items implies identical cardinals.

There are important methodological differences between the Frydman and Bryant, and Sophian et al., studies that may contribute to their divergent conclusions. First, Sophian et al.'s study started out with conceptually paired items; each frog had its own boat. Separating out six pairs of frogs and boats produces a set of six frogs and a set of six boats. In contrast, in Frydman and Bryant's task, sharing out six sweets gives two sets of three sweets. If conceptual pairs are thought of as whole (and not separate) entities, then the elements – the frogs and boats – become corresponding portions of these whole items. Making inferences about equivalence here need not be the same as making inferences when dividing up homogeneous items.

Another difference concerns the visibility of the sets. Children could see the set-to-be-inferred in the Frydman and Bryant but not the Sophian et al. (1995) study. In each case, children were essentially asked, "How many items are there in the first set?" followed by "How many items are there in the second set?" It may be that children count (rather than infer the numerosity of) a visible set simply because items are there. Children are prone to counting a single set in response to the question 'How many?' (Frye, Braisby, Lowe, Maroudas & Nicholls, 1989), particularly when items are visible (Fuson, 1988; Wynn, 1990). Number conservation tasks also demonstrate that the perceptual impact from visible arrays can influence, even if not determine, children's appreciation of the situation (e.g., Piaget, 1952). Furthermore, children may opt to repeat a strategy (counting) on the second question that proved successful on the first. The principal aim of the present studies was to investigate the influence of these two contextual factors - [1] sharing and [2] set-visibility - on children's developing ability to infer the number of items in a corresponding set.

We designed two play scenario's with which to examine the effects of these contextual variables. In one, that echoes Sophian et al.'s (1995) frog-boat task, children are shown toy animals that travel to school on their bicycles. Children are invited to pair animals with bicycles, before moving the pairs towards a schoolhouse. Upon arrival, the animals are placed inside the school, while the bicycles are placed in a separate bike shed. In the other scenario that echoes Frydman and Bryant's (1988) sweet-sharing task, children are asked to share out toy ice-creams between a toy ice-cream van and ice-cream shop. The questions put to children after allocating items is essentially the same in each scenario. Children are asked, 'How many animals are there in school/How many ice-creams are in the ice-cream shop?' followed by, 'How many bicycles are there in the bike shed/How many ice-creams are in the ice-cream shop?' On half the trials, the bike shed/ice-cream shop is covered, preventing children from seeing inside. If they are aware that corresponding sets have the same cardinal number, but prefer to count a target set rather than infer its numerosity, we expect children to make more inferences when the target set is hidden (hypothesis 1). If they know that sets of conceptually paired items (e.g., lions riding bicycles) have the same cardinal number but do not understand that sharing sweets evenly produces equivalent sets, then we expect children to perform better when they are quizzed about the number of lions (hypothesis 2).

We began to address these hypotheses with a relatively small-scale experiment. We restricted the number of trials to children, so as to minimize fatigue or carry-over effects, and consequently focused on a single set-size and age group.

Experiment 1

Method

Participants. Forty children (3 years 5 months to 4 years 3 months, *M* = 3 years 9 months, *SD* = 2.6 months), from three pre-school playgroups in a predominantly White, urban population took part. None of the children had any identified learning difficulties.

Materials. Child-friendly materials similar to children's toys were used throughout. Two shoe-boxes, painted on one side to resemble a bicycle shed and a school, and on the other side an ice-cream van and a seaside shop. Two sets of toy animals (lions and dolphins) and two sets of toy food items (ice-creams and tins of soup). Matchboxes painted to resemble bicycles.

Procedure. The same male experimenter tested each child on two separate occasions (Sessions 1 and 2), approximately one week apart. Problems were presented in two scenarios; animals going to school and food being allocated to vendors to sell. In each scenario, the problem was presented in two levels of visibility: visible and hidden. Each child received two trials, one in each level of visibility, in each session. The order of presentation was counterbalanced.

Scenario 1 - Animals going to school: Warm-up phase. The experimenter introduced the child to the school and bicycle shed, then demonstrated how the animals got to school each day, saying, "These lions go to school on their bikes. Look, each lion has his own bike, and he sits on it and pedals to school like this". Lions were placed in matchboxes (one per matchbox) and slid across the table, whereupon each lion was placed in the school (by inserting them through a small hole in the roof) and the matchboxes lined up outside the bicycle shed. Dolphins swam to school *en masse* (hidden in the experimenter's hand). In the warm-up phase, children were asked only which type of animals the bicycles belonged to.

Test phase. A set of lions (7, 8 or 9) was placed on the table, along with the same number matchboxes (bicycles). The experimenter asked the child to take the lions to school, reminding them that lions went one at a time. The experimenter moved the dolphins by hiding them in his hand and placing them inside the shoebox. The set of lions was the target set (i.e., the set in item-to-item correspondence with the bicycles) while dolphins acted as a control set (i.e., this set was not in one-to-one correspondence, and acted as a check against last-word-repetition using the cardinal number of bicycles). Once all the animals were in school the experimenter asked, "How many bicycles are in the bicycle shed?" followed by "How many lions/dolphins are at school?" (the order in which children were asked about the target and the control sets was counterbalanced). If they said they did not know, children were asked to guess. The aim was to see if children would use cardinal extension to infer appropriately the number of lions from the number of bicycles, and refrain from extending the same cardinal to an unknowable number of dolphins.

In the hidden condition the roof was left on ("because it is raining"), hiding the animals from view. In the visible condition the roof was removed ("because it is a nice day"), revealing the animals inside. If a child started to count the target set, the experimenter placed his hands over the shoebox and asked if they knew how many there were without counting. The bicycles were visible throughout both conditions. As an additional control against children succeeding through perceptual comparisons, all sets were arranged such that there was no spatial correspondence between any of them. Children readily rely on spatial cues to judge relative magnitude (e.g., Brainerd, 1979; Cowan, 1987; Michie, 1984a, 1984b; Piaget, 1952; Saxe, 1977; Sophian, 1987), and we wanted to eliminate the potential for children to make comparisons this way.

Scenario 2 - Food being sold: Warm-up phase. The child was asked to identify the ice-cream van and the shop, and then shown toy ice-creams. The experimenter then demonstrated how, as neither the shop nor the van had anything to sell (the shoeboxes were empty), it was necessary to share the ice-creams out so they both had the same. The child was invited to share them out after the experimenter had demonstrated how it could be done, and also shown some tins of soup (the experimenter explaining that when the weather got cold the shop needed to have soup in case people wanted something hot to eat).

Test phase. A set of ice-creams (14, 16 or 18) was placed on the table, and 3 - 5 tins of soup were retained in the experimenter's hand after being shown briefly to the child. The procedure was essentially the same as used for animals, with ice creams being allocated to the shop (the target set) and van (corresponding set) except that children were told, "Share out these ice-creams, so that the ice-cream shop and the ice-cream van have both got some ice-creams to sell. Put some in here [pointing] and some in here so they've both got the same." Those placed in the shop became the target set and those in the van the corresponding set. Tins of soup acted as the control (unrelated) set. Once all the ice creams had been shared out, the experimenter asked "How many ice creams does the ice cream van have to sell?" followed by "How many ice creams/tins of soup does the shop have to sell?" (the order in which the experimenter asked about the target and control sets was counterbalanced). Again, the corresponding set – ice-creams in the van - was visible on all trials.

Results

We start by addressing hypotheses 1 and 2 concerning the effects of set-visibility and scenario on the ability to infer corresponding numerosity. First we take as our dependent variable the ability to make an appropriate inference spontaneously; that is without first attempting to count the target set. Next, we recode performance (on trials where the target set was visible). Following recoding, the dependent variable will include all correct inferences (i.e., spontaneous inferences and the inferences made after the experimenter had hidden the target set with his hand because the child initially attempted to count the set).

For all analyses, a child was scored as correct when he or she inferred that the target (related) set, but not the unrelated set, shared the same cardinal as the counted set[1]. A preliminary analysis of the data revealed no effect of age (in months) or order of questioning (asking about the target/unrelated set first) and, consequently, the data were pooled for all subsequent analyses.

Spontaneous inferences. Table 1 shows the effect that set-visibility and scenario had on performance. Only 5% of children inferred spontaneously a visible set of ice creams, and 12.5% inferred a visible set of lions (see left hand column of Table 1). These figures contrast sharply with the inferences made when the ice-creams and lions were hidden (30% and 40% respectively - see right hand column of Table 1).

Table 1

Number of Children making inferences in each scenario

Visibility	Visible		Hidden
	Spontaneous (no hiding)	Inc. inferences made after hiding	
Scenario			
Lions/Bicycles (conceptual pairs)	5	11	16
Ice-creams (shared homogeneous items)	2	12	12

Notes.

Spontaneous inferences are those where the child inferred without making any attempt to count items. The figures under 'Inc. inferences made after hiding' refer to the total number of correct inferences made, including those made after the experimenter hid the target set because the child attempted to count.

The possible maximum in each cell = 40

Because the data failed to satisfy the criteria for parametric analysis, we used Sign tests[2] to examine the effects of Visibility (visible vs hidden) and Scenario (animals vs food items). Regarding visibility, 21 children scored equally on 'visible' and 'hidden' trials (e.g., they scored 0-0 respectively). Of the remaining 19 children, 17 made more inferences when the target set was hidden, with only 2 children showing the opposite pattern (i.e., making more correct inferences when the target set was visible); p = .001. Thus, children who knew that the numerosity of a related set could be inferred were more likely to display that knowledge when that set could not be counted. Turning to the effect of Scenario, 26 children scored equally on 'Lions/Bicycles' and 'Ice-cream' trials. Of the remaining children, 10 made more correct inferences when judging the number of bikes, with only 4 children making more correct inferences when asked about ice-creams. Although children were, therefore, more successful when judging paired rather than shared items, this difference failed to reach significance; $p = .180$.

We looked for reasons why only 35 correct inferences were made out of a possible total of 160. When the target set was visible, only seven correct inferences were made; on the remaining 'visible' trials children often attempted to count the target set (54/73). The incorrect responses in the 'hidden' condition were an approximately even mix of guesses (22 trials) and "Don't know" (30 trials). We expected some children to try and remove the lid from the box when asked about hidden sets and approximately half did so. When this happened, children were prevented from seeing inside and asked the question again.

Whilst we were interested primarily in whether children would [1] infer or [2] try to count the target set, there was at least one other perfectly reasonable strategy open to the children, which was to recount the set they had just counted. However, none of the children resorted to this approach. We also accepted attempts to count the unrelated set - when that set was visible - as constituting correct responses. This was considered reasonable because the cardinal number for the counted set offered no clues as to the number of items in a set that was not in one-to-one correspondence. It is permissible, on logical grounds, to count one set and infer the numerosity of a corresponding set. There are no grounds for extending the cardinal from the counted set to an unrelated set. Thus the unrelated sets of dolphins or tins of soup acted as a check on children's grasp of the logico-mathematical relationship between corresponding sets. Only two children successfully inferred the target set but failed to score because they made an incorrect inference about the unrelated set by extending the same cardinal to both sets.

One explanation for the pronounced influence of set-visibility might be that children were learning about the demands of the task on the earlier trials (i.e., on Session 1) and only later applying that knowledge. Although counterbalancing the order of presentation should have ruled this out, we checked by re-coding responses in terms of the chronological (first, second, third and fourth) order the children experienced the four trials. There were 15 correct inferences made on Session 1, and 20 made on Session 2: 8, 7, 9 and 11 on the first, second, third and fourth problems respectively. None of these differences were significant.

Inferences made after hiding the visible set. In the previous analyses, children were scored as correct only if they inferred spontaneously, without making any attempt to count the target set. Frydman and Bryant (1988) found that 10/24 children in their study did make the correct inference after the experimenter hid the target set and asked if the child knew how many there were without counting. In order to assess whether a similar effect occurred here, scores on trials in the visible condition were recoded to account for an inference whether it was spontaneous or not (i.e., a child was scored correct if they made an inference before or after the set had been hidden by the experimenter). The total scores on trials in the visible condition rose from 7 to 23 (see the middle column of Table 1). Comparing the number of correct inferences made on hidden-sets trials with the number made on the recoded visible-set trials, McNemar's test revealed there was no longer an effect of set-visibility; χ^2 $(1, N = 40) = .31$, $p > .05$. This reinforces the point that set-visibility was masking children's competence. One question is whether the increased scores following recoding were specific to a particular scenario. There were 27 correct inferences made when the target set consisted of lions inside the school and 24 when that set was the ice-creams inside the shop. Thus, following recoding, there was still no effect of scenario.

Although the small sample size in this first study precluded a statistical analysis of the roles that counting and sharing proficiency might play in children's ability to use cardinal extension, it is still interesting to note how well children were able to carry out these basic procedures. Whilst all children knew how to divide the paired sets of lions and bicycles by placing them in their appropriate locations, they were not as good at sharing the ice-creams out. Eleven children showed no ability to share evenly, 14 shared evenly on one trial, and 15 children distributed equal shares on both trials. Sharing is an emerging skill for this age group. As a measure of counting proficiency, we looked at the number of times children counted the corresponding sets (i.e., the sets the experimenter asked the child to count before asking about the target and control sets) correctly across the two scenarios. Twenty-two out of 40 were accurate on all four trials. A further nine children counted accurately on three of the four trials. Thus, these 31 children counted accurately on 115 out of 124 trials, displaying little evidence that they had much difficulty enumerating the set-sizes we used. The remaining nine children showed relatively poor counting skills, counting inaccurately on at least two of the four trials. This shows that counting, like sharing, was also an emerging skill for some of these children (Counting proficiency was significantly correlated with sharing proficiency; $r(39) = .41$, $p < .01$). However, given the small number of inferences made spontaneously (35/160), and the finding that most children were good at counting the sets we used, we find no grounds for suspecting that a failure to infer was due to a lack of self-confidence in their counting accuracy. We explore this issue further in Experiment 2.

Discussion

Nearly half the children we tested showed no understanding that numerical equivalence could be inferred using cardinal extension when the target set was in one-to-one correspondence with a set already counted. The biggest influence on whether children would display this ability was the visibility of the sets. Few children resisted counting when it was a strategy available to them; being able to see the target set appeared to encourage them to count. In contrast, hiding the target set seemed to promote numerical inferences. Clearly, for some children, a reluctance to infer a corresponding numerosity cannot be accounted for an absence of the requisite understanding. Children's knowledge may be underestimated if we take unnecessary counting as evidence that they have not yet grasped the relationship between correspondence and cardinal numbers.

We did not find reliable evidence that children were using the conceptual pairing of items as a cue to numerical correspondence, an association that is absent when homogeneous items are shared out. Both acts produce equivalent sets (when, as here, even numbers are used), and children did make more spontaneous inferences (albeit statistically non-significant) about paired sets than they did about sets that resulted from sharing, although this difference largely disappears when all correct inferences are considered. This is an issue we return to later.

Experiment 2

The results from Experiment 1 motivated a follow-up study that was more ambitious in scope, with the aim of replicating and extending the findings thus far. Hypotheses 1 and 2 were retained from Experiment 1 and re-assessed using more trials (and therefore, potentially greater power). An additional issue that we wanted to investigate concerned the sizes of the sets to be inferred. One possibility is that children in Experiment 1 were reluctant to make inferences about another set because they were sensitive to the limitations in their counting proficiency. We know that three-year-olds are able to infer equivalence when small sets are used (two / three items: Cowan & Biddle, 1989; Sophian et al., 1995). But small numerosities - certainly up to four items - can be enumerated non-verbally (Chi & Klahr, 1975; Starkey and Cooper, 1980; Trick & Pylyshyn, 1994), and can be compared without verbal counting (Strauss & Curtis, 1984; but see also Mix, Huttenlocher & Levine, 2002). The fact that young children judge more accurately when sets are small and visible raises the possibility that they are either making perceptual comparisons or enumerating the sets non-verbally. We therefore included small (two & three items) and larger sets (five & six items) in the stimuli. If children were less likely to make a numerical inference because they lack confidence in their counting ability (Gelman, 1972), we would expect larger sets to produce a different pattern of response (hypothesis 3). Determining children's counting proficiency permits an assessment of whether they can infer numerosity above their counting range. By including two different age groups (three- versus four-year-olds) we hope to build a developmental model of children's abilities by looking at the interaction between these factors.

Finally, we wanted to examine whether either of the two core skills of counting and sharing are associated with the ability to draw appropriate numerical inferences. Children begin to grasp cardinality between the ages of 3 and 4 (Fuson, Pergament, Lyons, & Hall, 1985; Wynn, 1990), around the same time as they are learning that sharing is a numerical activity (Miller, 1984). Research into whether sharing competence is related to counting skill has found no association between the two (Pepper & Hunting, 1998). However, Mix (1999a, 1999b) and Saxe (1977, 1979) found children's counting accuracy correlates with the recognition of numerical equivalence on set-comparison tasks. The ability to maintain one-to-one correspondence between words and items when counting appears to offer a developmental advance particularly where there are perceptual differences (e.g., in size, colour, etc.) between sets being compared. However, it is less clear how the ability to count single sets accurately might help children gain the conceptual insight that sharing evenly produces sets with the same cardinal. It is intuitively appealing to suspect that the ability to maintain item-to-item correspondence between shares is a more likely predictor of this ability.

Method

Participants. Nineteen three-year-olds (3 years 8 months to 4 years 0 months; *M* = 3 years 10 months, *SD* = 1.6 months) and 26 four-year-olds (4 years 1 month to 4 years 8 months; *M* = 4 years 4 months, *SD* = 2.2 months) from two pre-school playgroups in a predominantly White, urban population took part. None of the children had any identified learning difficulties. None had taken part in Experiment 1.

Materials. As in Experiment 1 except there were three sets of toy animals (lions, pandas and ladybirds). Ladybirds that 'flew' to school always acted as the control set (in place of dolphins) in Scenario 1, and lions and pandas were interchanged as the target set in an attempt to maintain children's interest in the tasks.

Procedure. The same as Experiment 1, except that in each scenario, the problems were presented with both small (2-3 items) and large (5-6 items) sets. Each child received 16 problems (8 in each Session), representing all the factorial combinations of the three factors. The order of presentation was counterbalanced using a Latin Square design. The control sets were only included on Session 2 in an attempt to reduce the time taken to administer the 16 trials and minimize the risk that children would withdraw from the study through fatigue or disinterest.

Results

Spontaneous inferences. There was no significant difference between the number of inferences made on Sessions 1 and 2 (182 versus 201), and so data were pooled across session. In the 'visible' condition, children tended to either infer spontaneously (141/360 trials) or attempt to count the target set (177/360), accounting for over 83% of responses in this condition. The remaining responses were a mix of guesses (27) and "Don't know" (15). When the target set was hidden, children made many more inferences spontaneously (242/360). The incorrect responses in the 'hidden' condition were, again, a mix of guesses (78) and "Don't know" (39), but no attempts were made to count the target set. This was a surprise, as we expected some children to try and remove the lid from the box to see inside, but none did. Table 2 shows the mean number of times children made a spontaneous and appropriate [3] inference about the items in the target set for each set-type and set-size.

The data satisfied the criteria for a parametric test, and ANOVA was therefore not only suitable, but also corresponds with the analyses conducted by Sophian et al. (1995). A 2 (Age: 3 vs. 4) x 2 (Scenario: animals vs. ice-creams) x 2 (Set size: small [2-3] vs. large [5-6]) x 2 (Visibility: hidden vs. visible) repeated measures ANOVA, with Scenario, Set size and Visibility as within-subjects variables, confirmed hypothesis 1 in that the main effect of Visibility was highly significant, F (1, 43) = 56.06, $p < .001$; η_p^2 (i.e., partial *Eta* squared) = .57.

Table 2

Mean number of correct spontaneous inferences and inferences made after hiding the visible set (max. possible 2 in each condition)

	Hidden		Visible	
			Spontaneous (after hiding)	
Set size	Small	Large	Small	Large
Age/Scenario				
3-year-olds				
Ice-creams				
M	84	.63	.37 (1.00**)	.32 (.58)
SD	.60	.89	.60 (.67)	.58 (.69)
Animals				
M	1.74	1.11	.74 (1.37**)	.47 (.95**)
SD	.56	.80	.87 (.68)	.70 (.78)
4-year-olds				
Ice-creams				
M	1.54	1.08	1.15 (1.50*)	.58 (1.00**)
SD	.58	.84	.83 (.51)	.76 (.69)
Animals				
M	1.81	1.73	1.31 (1.73*)	1.00 (1.27*)
SD	.49	.60	.84 (.45)	.80 (.67)

Note. Asterisks denote significant differences between the number of correct inferences made after the experimenter had hidden the relevant set and the number of spontaneous inferences given for those sets.

* $p < .05$ ** $p < .01$

For every combination of Scenario and Set size, children made more spontaneous inferences when the target set was hidden from view. A significant interaction between Scenario and Visibility, F (1 43) = 8.72, $p < .05$; $\eta_p^2 = .17$, showed that the influence of set visibility was greater on 'Animals' problems than 'Ice-cream' problems. There were also two significant 3-way interactions, both including Age.

There was an interaction between Age, Visibility and Set size, F (1, 43) = 4.25, $p < .05$, $\eta_p^2 = .09$. While all simple effects of Visibility were significant on both small and large sets in both age groups, for three-year-olds, the effect of Visibility was most marked on small sets [F (1, 17) = 32.67, $p < .001$, $\eta_p^2 = .65$ vs F (1, 17) = 9.92, $p < .01$, $\eta_p^2 = .35$], whereas for four -year-olds, the effect was most pronounced on large sets [F (1, 24) = 10.90, $p < .005$, $\eta_p^2 = .30$ vs F (1, 24) = 34.41, $p < .001$, $\eta_p^2 = .58$]. Three-year-olds are better at inferring small sets than large sets, especially when those sets are hidden. Conversely, four-year-olds are generally able to infer small sets regardless of whether they are hidden or not, but some have difficulty inferring larger sets.

A second 3-way interaction was found between Age, Set size and Scenario, F (1, 43) = 12.52, $p < .005$, $\eta_p^2 = .23$. All simple effects of Scenario were significant for both small and large sets for each age group. Thus, although the effect of scenario failed to reach significance in Experiment 1, the greater power in the design of Experiment 2 has revealed this effect to be a reliable one. For three-year-olds, the effect of Scenario was most marked on small sets [F (1, 17) = 21.46, $p < .001$, $\eta_p^2 = .54$ vs F (1, 17) = 5.59, $p < .05$, $\eta_p^2 = .24$], whereas for four -year-olds, the effect was most pronounced on large sets [F (1, 24) = 6.34, $p < .05$, $\eta_p^2 = .20$ vs F (1, 24) = 22.27, $p < .001$, $\eta_p^2 = .47$]. Three-year-olds are not as good at inferring equivalent shares as they are at recognizing that conceptually paired items share the same cardinal, even when smaller sets are used. In contrast, four-year-olds are almost as good at recognizing the relationship between sharing and cardinality as they are at recognizing the relationship between cardinality and conceptually paired items, although they have greater difficulty in grasping that larger shares also have identical cardinals.

The other significant results were a main effect of Age, $F\ (1, 43) = 9.96$, $p < .005$, $\eta_p^2 = .19$; a main effect of Set size, $F\ (1, 43) = 33.14$, $p < .001$, $\eta_p^2 = .44$; and a main effect of Scenario, $F\ (1, 43) = 38.23$, $p < .001$, $\eta_p^2 = .47$. In general, older children made more spontaneous inferences than younger children; the number of spontaneous inferences was greater when sets were small (supporting hypothesis 3); and children in both age groups were more likely to recognize the equivalence of paired items than they were to recognize the numerical equivalence that results from sharing things out evenly (supporting hypothesis 2).

We also investigated which, if either, of the core skills of counting and sharing, predicts the ability to infer the size of a corresponding set. Twenty-nine children (64%) in the sample were consistently able to maintain item-to-item correspondence when they shared small sets of 4 and 6 items, although only 20 were able to do so with larger sets of 10 and 12 items. This change was significant as determined by a McNemar χ^2 test for significance of change, $\chi^2=6.4$, $p< .05$. Sharing was an emerging skill for these children, and four-year-olds were significantly better than three-year-olds, $F\ (1, 43) = 4.16$, $p < .05$, $\eta_p^2 = .09$. Similarly, children were better at counting small sets than large sets, $F\ (1, 43) = 20.76$, $p < .001$, $\eta_p^2 = .33$, although three year olds were in general just as good as the older children, F (1, 43) = 1.38, $p = .25$, $\eta_p^2 = .03$. It seems that the ability to count sets of up to six items is in place by the age of three. Forty-one children (91% of the sample) were able to count sets of 2-3 items without any difficulty, and over half (56%) consistently counted larger sets accurately.

Hierarchical multiple regression was carried out using the number of spontaneous, correct inferences as the dependent variable[4]. Counting proficiency was measured as the number of trials [/16] where the child counted the corresponding set accurately (*M* = 14.78, *SD* = 2.27) and sharing proficiency as the number of trials [/8] where the child maintained item-to-item correspondence when sharing (*M* = 6.13, *SD* = 2.07).

Table 3

Hierarchical Multiple Regression for variables predicting correct inferences

Variable	□	*SE* □	□	□R^2
Spontaneous correct inferences[1]				
Step 1				
Age	2.31	1.04	.25*	.19**
Step 2				
Counting proficiency	0.38	.24	.19	.14**
Step 3				
Sharing proficiency	1.10	.27	.50**	.19**
All correct inferences[2] (including those made after hiding-the set)				
Step 1				
Age	1.70	.82	.19	.17**
Step 2				
Counting proficiency	0.65	.19	.33*	.28**
Step 3				
Sharing proficiency	1.16	.21	.54**	.23**

[1] R^2 = .52

[2] R^2 = .68

* $p < .05$

** $p < .01$

Given that the previous analyses had revealed a significant main effect of age, this variable was entered first (see Table 3), and was confirmed as a significant contributor to the explained variance; r (44) = .43. Although the zero-order correlations between spontaneous inferences and both counting proficiency and sharing proficiency were significant (.45 and .66, respectively), we entered counting proficiency next, followed by sharing proficiency. Together, these three variables explained over 50% of the total variance (see Table 3). Moreover, the model shows that sharing proficiency is a unique predictor above and beyond the influence of age and counting proficiency (In contrast, counting proficiency failed to exert a similar effect once sharing proficiency had been entered; $\Delta R^2 = .30$, $p > .05$).

Discriminating between the target and control sets. It is important to check that children were using cardinal extension to infer number because they knew that sets were equivalent and not simply using a last-word rule. Children could have made a correct inference about the target set by simply repeating the cardinal number for the counted set without grasping why the two sets had the same cardinal number. We recoded the data from Session 2 such that a child was scored as correct only if he or she also understood that the target set and control set had different cardinal values. If children were basing their judgments on quantitative correspondence they should be more likely to use cardinal extension for that set than for another set they had no numerical knowledge of.

There were 201 correct spontaneous inferences made during Session 2. On 182 of these trials (91%), children discriminated between the related target set and the control set. There were only 26 instances of children extending the cardinal from the counted set to both the target and control sets, with three-year-olds being more likely to make this mistake than four-year-olds, F (1, 43) = 4.52, $p < .05$, $\eta_p^2 = .10$. Other than the effect of Age, there were no other significant influences on children extending the cardinal inappropriately; Set size, Scenario and Visibility had no effect in this regard (all ps > .05). The control questions implied that only two children did not understand that the same cardinals apply only to equivalent sets. Thus, performance confirms that nearly all the children who used cardinal extension did so because they understood the relationship between cardinality and quantitative equivalence.

Inferences made after hiding a previously visible set. As in Experiment 1, scores on trials in the visible condition were recoded to discriminate between inferences made spontaneously and those made after the experimenter hid the items and repeated the question because the child went to count them. The relevant data are shown in Table 2. Hiding-then-repeating the question increased the number of inferences for almost every combination of Age, Scenario, Set size and Visibility. Overall, the number of valid inferences made rose from a total of 383 to a total of 488. This is striking given that children went to count the visible target set on only145 trials. Thus, when they were prevented from counting, children went on to infer correctly over 70% of the time. The exception was when three-year-olds were asked about large sets of ice-creams, and this led to a 4-way interaction between Scenario, Set size, Visibility and Age, F (1, 43) = 5.10, $p < .05$, $\eta_p^2 = .11$. Even when counting was prevented, by hiding the set, three-year-olds were relatively poor at inferring large sets of shared items relative to small sets or paired sets. The interactions between Scenario and Visibility, F (1, 43) = 7.22, $p < .05$, $\eta_p^2 = .14$, and between Scenario, Set size and Age, F (1, 43) = 4.88, $p < .05$, $\eta_p^2 = .10$, are both present as in the earlier analysis with similar estimates of effect sizes. We are not clear why hiding the target sets (when children started to count them) reduced, but did not eradicate, the effect of set-visibility. It is possible that some children, when denied their first-choice strategy, fail to search for an alternative, although we concede that our data cannot shed any light on this matter.

Recoding the data also retains all four significant main effects of Set size, Age, Scenario, and Visibility (all ps < .01, all $\eta_p^2 > .08$). The estimates of effect size are similar to before except that a smaller effect of Visibility ($\eta_p^2 = .09$ vs .57) confirms that being able to see a target set encouraged unnecessary counting. Attempting to count the target set was not associated with counting proficiency: $r(44) = .13, p > .05$. Furthermore, as Table 3 shows, it is still sharing proficiency that makes the biggest unique contribution to the model (r [44] = .74) over and above the contributions of counting proficiency (r [44] = .60) and age (r [44] = .41).

In sum, the various analyses support the notion that important precursor to recognizing the link between corresponding sets and cardinal numbers are age-related developments in sharing and counting proficiency. Improvements in each of these basic skills are likely to be associated with the ability to draw appropriate inferences about numerical relationships, but this understanding often remains hidden when alternative strategies like counting are available or seen to be encouraged.

Discussion

As in Experiment 1, the visibility of arrays appears to act as a cue to counting. However, the introduction of an extra age group revealed some important differences. In contrast to Experiment 1, there was now much stronger evidence that the youngest children used conceptual pairing as a clue to numerical correspondence, and that the absence of this bond when homogeneous sets are shared out leads to something of an arithmetical *impasse* for three-year-olds. Both acts produce equivalent sets, but younger children were more likely to infer equivalence if there was an established association between items. It is interesting to compare this finding with other evidence that suggests a pattern of gradual de-contextualization in children's recognition of numerical correspondence. In set-matching tasks the degree of surface similarity (e.g., colour, shape) is important; three-year-olds recognize equivalence between homogeneous sets (e.g., one set of dots and another of shells) before they are able to recognize the relationship between sets of mixed objects around the age of four (Mix, 1999a). The level of perceptual support together with mastery of the counting routine helps very young children to make quantitative comparisons, at least when those sets are small (fewer than five items). Notwithstanding this association, it remains unclear how the ability to enumerate single sets helps children to grasp the fact that cardinal numbers have relative value.

Indeed, improvements in counting accuracy do not always correlate with the use of counting strategies to make numerical comparisons (Saxe, 1979). Moreover, counting proficiency appears insufficient to predict children's understanding of number conservation (Piaget, 1952), indicating that even accurate enumerators still have something fundamental to learn about the significance of cardinal numbers. Procedural mastery must be supplemented by conceptual insight if children are to apply their understanding of cardinality to simple arithmetical problems like set-comparison (Muldoon, Lewis & Freeman, 2003). There are grounds for believing that conceptual knowledge of a procedure like counting, and its application as an arithmetical tool are linked (Sophian, 1997). Children need to learn the implications of cardinal numbers for judgments about relative quantity, not simply the connection between word-to-item correspondence and the cardinality of a single set (i.e., the one-to-one correspondence principle - Gelman & Gallistel, 1978), if they are to recognize why counting is a useful arithmetical tool.

Children appear to progress from being able to infer small conceptually related pairs of items before grasping the numerical significance of sharing, and this develops in tandem with age-related mastery of the sharing procedure. The use of small conceptually paired sets of items proved the least difficult of all our tasks and most children were able to infer appropriately. It is possible that conceptually paired sets, even pairs that are not found outside of children's games, as ours were, help children to access the fact that every item of one type has a corresponding item of another type. Items were 'paired' before being separated out. This is not usually the case when items are shared out. When sharing, set-equivalence is established for the first time following the sharing routine, and item-to-item correspondence is solely the result of temporal correspondence between distributions.

We remain unsure whether the effect of set-type is due to the 'pairing' procedure or the conceptual link between items. This could be examined by adding a third scenario to the two we employed here. This would entail 'sharing' paired items out without establishing the conceptual link beforehand (i.e., without asking children to put lions and bicycles together before separating them out). Putting one lion in the school, and then one bicycle in the bike shed, followed by a lion in school and another bicycle in the shed and so on, would mirror the sharing routine but would result in sets of different items. Comparing performance on this scenario with the present lions/bicycles scenario would provide the effect of the original pairing procedure. Also, contrasting the inferences about 'shared' lion/bicycles in the new scenario with inferences about shared ice-creams would test the possibility that the greater success we observed with conceptually paired items was a result of them evoking a stronger 'equivalence' schema.

Our findings also suggest why the children in Frydman and Bryant (1988) failed to show the level of success achieved by younger children in Sophian et al. (1995). The success rate of 77% in Sophian's task was achieved only when sets were small (two/three items). When sets were larger – five/six items - performance dropped to less than 50%. It is possible that the greater temporal demand of sharing large numbers of items prevents children from monitoring their own maintenance of item-to-item correspondence, leading to a decrease in confidence in their own sharing accuracy. The finding that age-related change in sharing proficiency was the best predictor of cardinal extension highlights the impact that an emerging mastery of temporal item-to-item correspondence has on children's awareness of sharing as a numerical activity. It is presumably easier to remember that each recipient has received the same if only 4 items are distributed than if 12 items are shared out. The fact that children made more correct inferences (albeit not to a statistically significant degree) with smaller shares supports this explanation.

The finding that children use an effective strategy one moment only to discard it the next is not a new one. Siegler (1995, 1996) offers a different explanation. Variability in performance is a characteristic of a transitional phase in the development from one state of knowledge to another. Knowledge of different strategies, whether explicit or implicit, has been viewed as a marker of developmental readiness. A child might use different strategies to solve two examples of the same mathematical problem on successive days (Siegler & Shrager, 1984). If children use ineffective strategies when they have effective strategies at their disposal, development is a shift in the frequency with which the competing strategies are selected for use.

General Discussion

An important mathematical achievement for children is to recognize both [1] that corresponding sets have the same cardinal number, and [2] that sharing evenly produces such sets. It is clear that the recognition of numerical equivalence for conceptually paired items precedes the understanding that sharing also produces similar correspondence. Whilst three-year-olds might be aware that sharing produces equivalent sets (Cowan & Biddle, 1989), and recognize the relevance of numerical values for correspondence relations (Becker, 1989; Sophian, 1988), our results provide further evidence that they have difficulty in connecting the two. Procedural mastery of counting is typically in place by the age of three, but age-related improvements in the ability to maintain item-to-item correspondence in a temporal action predict children's developing understanding of the connection between sharing, set equivalence, and cardinality.

Of those who could make an appropriate inference, the visibility of the sets influenced whether they would display this understanding spontaneously. Many children were unable to resist counting when it was a strategy available to them. It is perhaps not surprising that being able to see a target set appears to encourage children to count (i.e., a familiar, 'direct' strategy) rather than infer (an 'indirect' strategy). Alternatively, it might be concluded that children tried to count items because they were not confident in the equivalence between the sets. We are less convinced by this possibility because having the target set hidden from the outset often revealed their confidence; the number of spontaneous inferences made when sets were hidden was over 70% higher than when sets were visible. It is worth noting here that these inferences about hidden sets were not tentative guesses; there was rarely, if ever, any suggestion that children were not confident in their judgments (even in the minority of cases when those judgments turned out to be wrong). It seems that some children possess a nascent concept of equivalence but circumstances can mask this. Consequently, children's knowledge may be underestimated if we take unnecessary counting as evidence that they have yet to grasp the link between quantitative correspondence and cardinal numbers. Children's abilities may be revealed one moment but remain unexpressed the next if they do not feel it necessary to apply those particular skills.

There are at least five possible explanations why children might persevere with counting even when they are able to infer. First, children typically count when asked 'How many?' (e.g., Frye et al., 1989; Fuson, 1988; Wynn, 1990). Secondly, children may be encouraged to count by being asked 'How many?' for both sets; the effect of asking the same question twice has been clearly demonstrated on number conservation tasks (e.g., Rose & Blank, 1974; Samuel & Bryant, 1984). Thirdly, children may believe that counting is not merely a check, but essential if one is to know whether shares are fair (Davis & Pitkethly, 1990; Desforges & Desforges, 1980). Selective training where children are asked to make numerical inferences about hidden arrays might reduce this type of perseveration. Using small sets of two and three items offers the potential for even three-year-olds to begin abstracting cardinal numbers in a way that perceptual cues to quantity often inhibit.

It is also possible, however, that performance was influenced by other domain-general information processing constraints. Working memory is needed for the active maintenance of temporary representations (Baddeley, 1986), such as the cardinal for the corresponding set in the present studies. Within-task forgetting is likely to impact on children's performance on mathematical tasks, and there is evidence suggests that they employ a task-switching strategy when carrying out a counting span task that incorporates memory and processing requirements (Towse & Hitch, 1995; Towse, Hitch & Hutton, 1998). That is, they engage either in memory operations, or counting operations, but not both simultaneously. Thus a fourth explanation is that children simply forgot information encoded in memory when asked an additional question about a different set.

A fifth explanation is that the development of executive skills plays an important role. The results from Experiment 2 show that the ability to judge numerical equivalence using cardinal extension develops sometime between the ages of three and four. A concomitant growth in children's reflection on their own behaviour between these ages increases the amount of control they exert over their actions, but younger children are prone to ignore rules, even when they are able to verbally demonstrate appropriate rule-knowledge (Zelazo, Frye & Rapus, 1996). Thus executive limitations might account for the fact that many children in the present studies apparently knew the rule 'Extend the cardinal from the corresponding set to the target set' but did not apply it across all trials. Further work is needed to resolve which, if any, of these alternative explanations is valid.

To summarize, we began by noting that children's numerical processing is often characterized by a perplexing mix of abilities and limitations (Feigenson et al., 2004). The present datasets show that, even within a particular paradigm (that is, with formally comparable tasks), there can be substantial differences in what children do. However, the data also help us to understand the basis of such variability. We argue that children can make inferences about numerical correspondence, but the physical presence of each set draws children towards counting in any case, and the set array contributes to this tendency too. We also argue, in line with others (Davis & Pitkethly, 1990; Desforges & Desforges, 1980; Frydman and Bryant, 1988) that three-year-olds have yet to develop a full appreciation of the connection between sharing and cardinality. They are in a transitional stage, in which the concept of one-to-one correspondence, although central to mathematical reasoning and an understanding of what number represents (e.g., Piaget, 1952; Russell, 1960), is fledgling in character. Although at this stage little is known about how children's understanding of sharing develops, we suggest that improvements in sharing proficiency are likely to be associated with a greater understanding of sharing as a quantitative exercise. Moreover, presenting children with arithmetical situations that preclude set enumeration may be especially helpful in directing them towards the conceptual realization that counting “because it’s there” is unnecessary.

References

Baddeley, A. (1986). Working Memory. Oxford: OUP.

Becker, J. (1989). Preschoolers' use of number words to denote one-to-one correspondence. Child Development, 60, 1147 - 1157.

Brainerd, C. (1979). The origins of the number concept. New York: Praeger.

Chi, M. T. H., & Klahr, D. (1975). Span and rate of apprehension in children and adults. Journal of Experimental Child Psychology, 19, 157 - 192.

Cowan, R. (1987). When do children trust counting as a basis for relative number judgments? Journal of Experimental Child Psychology, 43, 328 - 345.

Cowan, R., & Biddle, S. (1989). Children's understanding of one-to-one correspondence in the context of sharing. Educational Psychology, 9, 133 - 140.

Davis, G. E., & Pitkethly, A. (1990). Cognitive aspects of sharing. Journal for Research in Mathematics Education, 21, 145 - 153.

Desforges, A., & Desforges, C. (1980). Number-based strategies of sharing in young children. Educational Studies, 6, 97 - 109.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. Trends in Cognitive Sciences, 8, 307-314.

Francis, B., Green, M., & Payne, C. (Eds.). (1993). The GLIM System Release 4 Manual. Oxford: Oxford University Press.

Frydman, O., & Bryant, P. (1988). Sharing and understanding of number equivalence by young children. Cognitive Development, 3, 323 - 339.

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. Child Development, 60, 1158 - 1171.

Fuson, K. C. (1988). Children's counting and concepts of number. New York: Springer-Verlag.

Fuson, K. C., Pergament, G. G., Lyons, B. G., & Hall, J. W. (1985). Children's conformity to the cardinality rule as a function of set size and counting accuracy. Child Development, 56, 1429 - 1436.

Gelman, R. (1972). The nature and development of early number concepts. In H. W. Reece (Ed.), Advances in Child Development and Behavior (Vol. 7, pp. 115 – 167). New York: Academic Press.

Gelman, R., & Gallistel, C. R. (1978). The child's understanding of number. Cambridge, MA: Harvard University Press.

Gelman, R., & Meck, E. (1983). Preschoolers' counting: Principles before skill. Cognition, 13, 343-359.

Michie, S. (1984a). Number understanding in preschool children. British Journal of Educational Psychology, 54, 245 - 253.

Michie, S. (1984b). Why preschoolers are reluctant to count spontaneously. British Journal of Developmental Psychology, 2, 347 - 358.

Miller, K. (1984). Child as the measurer of all things: measurement procedures and the development of quantitative concepts. In C. Sophian (Ed.), Origins of Cognitive Skills (pp 193-228). Hillsdale, NJ:Erlbaum.

Mix, K. S. (1999a). Similarity and numerical equivalence: Appearances count. Cognitive Development, 14, 269 - 297.

Mix, K. S. (1999b). Preschoolers' recognition of numerical equivalence: Sequential sets. Journal of Experimental Child Psychology, 74, 309 - 332.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (1996). Do preschool children recognize auditory-visual numerical correspondences? Child Development, 67, 1592 - 1608.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Quantitative development in infancy and early childhood. Oxford: Oxford University Press.

Muldoon, K., Lewis, C., & Freeman, N. H. (2003). Putting counting to work: preschoolers' understanding of cardinal extension. International Journal of Educational Research, 39, 695 - 718.

Pepper, K. L., & Hunting, R. P. (1998). Preschoolers' counting and sharing. Journal for Research in Mathematics Education 29, 164 - 183.

Piaget, J. (1952). The child's conception of number. London: Routledge and Kegan Paul.

Rose, S. A., & Blank, M. (1974). The potency of context in children's cognition: an illustration through conservation. Child Development, 45, 499 - 502.

Russac, R. J. (1978). The relation between two strategies of cardinal number: Correspondence and counting. Child Development, 49, 728 – 735.

Russell, B. (1960). Introduction to mathematical philosophy. London: Allen & Unwin.

Samuel, J., & Bryant, P. (1984). Asking only one question in the conservation experiment. Journal of Child Psychology and Psychiatry, 25, 315 - 318.

Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: preschoolers see number words as specific. Cognition, 92, 329-352.

Saxe, G. B. (1977). A developmental analysis of notational counting. Child Development, 48, 1512 - 1520.

Saxe, G. B. (1979). Developmental relations between notational counting and number conservation. Child Development, 50, 180 - 187.

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. Cognitive Psychology, 28, 225 – 273.

Siegler, R. S. (1996). Emerging minds: The process of change in children’s thinking. Oxford: Oxford University Press.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), The origins of cognitive skills (pp. 229 – 293). Hillsdale NJ: Erlbaum.

Sophian, C. (1987). Early developments in children's use of counting to solve quantitative problems. Cognition and Instruction, 4, 61 - 90.

Sophian, C. (1988). Early developments in children's understanding of number: Inferences about numerosity and one-to-one correspondence. Child Development, 59, 1397 - 1414.

Sophian, C., Wood, A. M., & Vong, K. I. (1995). Making numbers count: The early development of numerical inferences. Developmental Psychology, 31, 263 - 273.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. Science, 210, 1033 - 1035.

Strauss, M. S., & Curtis, L. E. (1984). Development of numerical concepts in infancy. In C. Sophian (ed.), Origins of Cognitive Skills, (pp 131-155). Hillsdale, NJ: Erlbaum.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. Psychological Review, 101, 80 - 102.

Towse, J. N., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? Quarterly Journal of Experimental Psychology, 48A, 108 – 124.

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. Journal of Memory and Language, 39, 195 - 217.

Wynn, K. (1990). Children's understanding of counting. Cognition, 36, 155 - 193.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month olds infants. Cognition, 74, B1-B11.

[1] Children were scored as correct if their inferred number matched the cardinal for the corresponding set irrespective of whether that set had been counted accurately or not. Children's grasp of cardinality can be masked by procedural errors. For example, if the corresponding set contained 5 items but the child counted 6, he or she was scored correct if they inferred the target set to be 6. Three children scored on this basis on one occasion each.

[2] The Sign test compares the number of positive and negative differences between two repeated measures. For example, comparing the number of children who are better at inferring hidden sets (e.g., positive difference) with the number of children who are better at inferring visible sets (e.g., negative difference), where the Null hypothesis is that the number of positive and negative differences will be the same.

[3] Children were scored as correct if their inferred number matched the cardinal for the corresponding set irrespective of whether that set had been counted accurately or not. Children's grasp of cardinality can be masked by procedural errors. For example, if the corresponding set contained 5 items but the child counted 6, he or she was scored correct if they inferred the target set to be 6. Three children scored on this basis on one occasion each.

[4] We combined performance on both scenarios into separate dependent variables for [1] spontaneous inferences and [2] all inferences (i.e., including those made after the target set had been hidden). Tests revealed a robust degree of reliability in each case; [1] = .89, [2] = .90.